ECE 509: Convex Optimization **Rutgers: Spring 2024**

Homework #7

April 11, 2024

Name: Ravi Raghavan *Extension: No*

Problem 1. *Voronoi description of halfspace.* Let a and b be distinct points in $\mathbb{R}^n$. Show that the set of all points that are closer (in Euclidean norm) to a than b, *i.e.*, $\{x : ||x-a||_2 \leq ||x-b||_2\}$, is a halfspace. Describe it explicitly as an inequality of the form $c^T x \leq d$. Draw a picture.

Solution. We can see that the following two sets are equivalent:

$S_1 = S_2$ where $S_1 = \{x : ||x-a||_2 \leq ||x-b||_2\}$ and $S_2 = \{x : ||x-a||_2^2 \leq ||x-b||_2^2\}$

Let's work with S_2 since it will be a lot easier

$S_2 = \{x : ||x-a||_2^2 \leq ||x-b||_2^2\}$
$S_2 = \{x : ||x||_2^2 - 2 < x, a > + ||a||_2^2 \leq ||x||_2^2 - 2 < x, b > + ||b||_2^2\}$
$S_2 = \{x : -2 < x, a > + ||a||_2^2 \leq -2 < x, b > + ||b||_2^2\}$
$S_2 = \{x : 2 < x, b-a > \ \leq ||b||_2^2 - ||a||_2^2\}$
$S_2 = \{x : 2(b-a)^T x \ \leq ||b||_2^2 - ||a||_2^2\}$
$S_2 = \{x : (b-a)^T x \ \leq 0.5(||b||_2^2 - ||a||_2^2)\}$

This is a closed half-space

Problem 2. Which of the following sets S are polyhedra? If possible, express S in the form $S = \{x | Ax \preceq b, Fx = g\}$

(b) Yes S is a polyhedra.

Let $M_1 = [a_1, a_2, ..., a_n] \in \mathbb{R}^{1xn}$ and let $M_2 = [a_1^2, a_2^2, ..., a_n^2] \in \mathbb{R}^{1xn}$

Let F be the vertical concatentation of 1^T, M_1, and M_2. Let g be $[1, b_1, b_2]^T$

Let $A = -I$

We can express S, via compact notation, as $S = \{x | Ax \preceq 0, Fx = g\}$

(c) S is NOT a Polyhedra

Problem 3. *Hyperbolic sets.* Show that the *hyperbolic set* is $\{x \in \mathbb{R}_+^2 : x_1 x_2 \geq 1\}$ is convex. As a generalization, show that $\{x \in \mathbb{R}_+^2 : \prod_{i=1}^n x_i \geq 1\}$ is convex. *Hint.* If $a, b \geq 0$ and $0 \leq \theta \leq 1$, then $a^\theta b^{1-\theta} \leq \theta a + (1-\theta) b$

Solution. Let S be $\{x \in \mathbb{R}_+^2 : \prod_{i=1}^n x_i \geq 1\}$. Let's have two vectors j and k that are in S. Let the elements of j be $j_1, j_2, ..., j_n$. Let the elements of k be $k_1, k_2, ..., k_n$

We want to prove that $\theta j + (1-\theta) k \in S$.

Since we know that $j \in S$ and $k \in S$, we can state the following:

- $j_1 \geq 0, j_2 \geq 0, ..., j_n \geq 0$
- $j_1 j_2 j_3 ... j_n \geq 1$
- $k_1 \geq 0, k_2 \geq 0, ..., k_n \geq 0$
- $k_1 k_2 k_3 ... k_n \geq 1$

For $i \in [1, n]$, $j_i, k_i \geq 0$ and $0 \leq \theta \leq 1$, we can see that $0 \leq j_i^{\theta} k_i^{1-\theta} \leq \theta j_i + (1-\theta) k_i$.

We can also see that:

$\prod_{i=1}^{n} j_i^{\theta} k_i^{1-\theta} = (\prod_{i=1}^{n} j_i)^{\theta} (\prod_{i=1}^{n} k_i)^{1-\theta}$

Since $j_1 j_2 j_3 ... j_n \geq 1$ and $k_1 k_2 k_3 ... k_n \geq 1$, we can say that:

$\prod_{i=1}^{n} j_i^{\theta} k_i^{1-\theta} = (\prod_{i=1}^{n} j_i)^{\theta} (\prod_{i=1}^{n} k_i)^{1-\theta} \geq 1$

Since $j_i^{\theta} k_i^{1-\theta} \leq \theta j_i + (1-\theta) k_i$,

$1 \leq \prod_{i=1}^{n} j_i^{\theta} k_i^{1-\theta} \leq \prod_{i=1}^{n} \theta j_i + (1-\theta) k_i$

We have shown that $\theta j_i + (1-\theta) k_i \in S$ and that S is a convex set!

Since we have proved the generalized case, we can say that $\{x \in \mathbb{R}_+^2 : x_1 x_2 \geq 1\}$ is convex as well

---

Problem 4. Problem 2.16:
Show that if S_1 and S_2 are convex sets in $\mathbb{R}^{m+n}$, then so is their partial sum
$S = \{(x, y_1 + y_2) : x \in \mathbb{R}^m, y_1, y_2 \in \mathbb{R}^n, (x, y_1) \in S_1, (x, y_2) \in S_2\}$

Solution. Let's say that we have two points in S, namely $(x_1, y_{11} + y_{12})$ and $(x_2, y_{21} + y_{22})$. To prove that S is convex, we need to show that $\theta(x_1, y_{11} + y_{12}) + (1-\theta)(x_2, y_{21} + y_{22})$ is in S.

Based on the definition of S, we can see the following:

- $(x_1, y_{11}) \in S_1$
- $(x_1, y_{12}) \in S_2$
- $(x_2, y_{21}) \in S_1$
- $(x_2, y_{22}) \in S_2$

Since S_1 and S_2 are convex, based on the definition of convex sets, we can see that:

- $\theta(x_1, y_{11}) + (1-\theta)(x_2, y_{21}) \in S_1$
  $(\theta x_1 + (1-\theta) x_2, \theta y_{11} + (1-\theta) y_{21}) \in S_1$
- $\theta(x_1, y_{12}) + (1-\theta)(x_2, y_{22}) \in S_2$
  $(\theta x_1 + (1-\theta) x_2, \theta y_{12} + (1-\theta) y_{22}) \in S_1$

By definition of Set S, we can see that:

$(\theta x_1 + (1-\theta)x_2, \theta y_{11} + (1-\theta)y_{21} + \theta y_{12} + (1-\theta)y_{22}) \in S$

$(\theta x_1 + (1-\theta)x_2, \theta(y_{11} + y_{12}) + (1-\theta)(y_{21} + y_{22})) \in S$

$\theta(x_1, y_{11} + y_{12}) + (1-\theta)(x_2, y_{21} + y_{22})$ is in S.

Problem 5 (Problem 2.19(a)). *Linear-fractional functions and convex sets.* Let $f : \mathbb{R}^m \to \mathbb{R}^n$ be the linear-fractional function

$$f(x) = (Ax + b)/(c^T x + d), \;\; dom f = \{x | c^T x + d > 0\} \tag{1}$$

In this problem, we study the inverse image of a convex set C under f, *i.e.,*

$$f^{-1}(C) = \{x \in dom f : f(x) \in C\} \tag{2}$$

For each of the following sets $C \subseteq \mathbb{R}^n$, give a simple description of $f^{-1}(C)$

Solution. Let's look at the halfspace $C = \{y : g^T y \leq h\}$(with $g \neq 0$).

$f^{-1}(C) = \{x : g^T((Ax + b)/(c^T x + d)) \leq h, c^T x + d > 0\}$

Since $c^T x + d > 0$
$f^{-1}(C) = \{x : g^T(Ax + b) \leq h(c^T x + d), c^T x + d > 0\}$
$f^{-1}(C) = \{x : g^T Ax + g^T b \leq hc^T x + hd, c^T x + d > 0\}$
$f^{-1}(C) = \{x : (g^T A - hc^T)x \leq hd - g^T b, c^T x + d > 0\}$
$f^{-1}(C) = \{x : (A^T g - ch^T)^T x \leq hd - g^T b, c^T x + d > 0\}$

Let's call a new vector $p^T = (A^T g - ch^T)^T$ and $q = hd - g^T b$
$f^{-1}(C) = \{x : p^T x \leq q, c^T x + d > 0\}$

We can see that $f^{-1}(C)$ is just the intersection of a halfspace and the domain of f!

Problem 6 (Problem 3.17). Suppose $p < 1, p \neq 0$. Show that the function

$$f(x) = (\sum_{i=1}^{n} x_i^p)^{\frac{1}{p}} \tag{3}$$

with $dom f = \mathbb{R}^n_{++}$ is concave. This includes as special cases $f(x) = (\sum_{i=1}^n x_i^2)^{\frac{1}{2}}$ and the *harmonic mean* $f(x) = (\sum_{i=1}^n \frac{1}{x_i})^{-1}$. *Hint.* Adapt the proofs for the log-sum-exp function and the geometric mean in 3.1.5

Solution. Gradient: $\frac{\partial f}{\partial x_k} = (\sum_{i=1}^n x_i^p)^{\frac{1}{p}-1} x_k^{p-1}$

Jacobian:
$\frac{\partial^2 f}{\partial^2 x_k} = (1-p)(\sum_{i=1}^n x_i^p)^{\frac{1}{p}-2}[x_k^{2p-2} - x_k^{p-2} \sum_{i=1}^n x_i^p]$
$\frac{\partial^2 f}{\partial x_k \partial x_l} = (1-p)(\sum_{i=1}^n x_i^p)^{\frac{1}{p}-2} x_l^{p-1} x_k^{p-1}$

It is clear that we can express $\nabla^2 f(x)$ as follows:
$\nabla^2 f(x) = (1-p)(\sum_{i=1}^n x_i^p)^{\frac{1}{p}-2}[qq^T - (\sum_{i=1}^n x_i^p)diag(x_1^{p-2}, x_2^{p-2}, \ldots, x_n^{p-2})]$ where $q_i = x_i^{p-1}$

For concavity, we need $v^T \nabla^2 f(x) v \leq 0$ to hold true for all $x \in \mathbb{R}^n_{++}$

$v^T \nabla^2 f(x) v = (1-p)(\sum_{i=1}^n x_i^p)^{\frac{1}{p}-2}[v^T qq^T v - (\sum_{i=1}^n x_i^p) v^T diag(x_1^{p-2}, x_2^{p-2}, \ldots, x_n^{p-2})v]$

$v^T \nabla^2 f(x) v = (1-p)(\sum_{i=1}^n x_i^p)^{\frac{1}{p}-2}[(\sum_{i=1}^n q_i v_i)^2 - \sum_{i=1}^n x_i^p (\sum_{i=1}^n x_i^{p-2} v_i)]$

$v^T \nabla^2 f(x) v = (1-p)(\sum_{i=1}^n x_i^p)^{\frac{1}{p}-2}[(\sum_{i=1}^n x_i x_i^{p-2} v_i)^2 - \sum_{i=1}^n x_i^p (\sum_{i=1}^n x_i^{p-2} v_i)]$

The Cauchy Schwartz Inequality tell us that $(a^T a)(b^T b) \geq (a^T b)^2$
If we set $a_i = x_i^{\frac{p}{2}}$ and we set $b_i = x_i^{\frac{p-2}{2}} v_i$, substituting this into the Cauchy Schwartz inequality allows us to see that

$(\sum_{i=1}^n x_i^p)(\sum_{i=1}^n x_i^{p-2} v_i) \geq (\sum_{i=1}^n x_i^{p-1} v_i)^2 = (\sum_{i=1}^n q_i v_i)^2$

By extension, we can see that
$(\sum_{i=1}^n x_i^p) v^T diag(x_1^{p-2}, x_2^{p-2}, \ldots, x_n^{p-2})v \geq v^T qq^T v$

$[v^T qq^T v - (\sum_{i=1}^n x_i^p) v^T diag(x_1^{p-2}, x_2^{p-2}, \ldots, x_n^{p-2})v] \leq 0$

Since $p < 1$ and $x_i \in \mathbb{R}_{++}$, $(1-p)(\sum_{i=1}^n x_i^p)^{\frac{1}{p}-2}$ is a positive constant. Hence, $v^T \nabla^2 f(x) v \leq 0$ which means that we have proven concavity!

Problem 7. Adapt the proof of concavity of the log-determinant function in §3.1.5 to show the following

(a) $f(X) = tr(X^{-1})$ is convex on $dom f = \mathbb{S}^n_{++}$

Solution. To verify convexity, we can consider an arbitrary line given by $X = Z + tV$ where $Z, V \in \mathbb{S}^n_{++}$
Let $\lambda_{v1}, \lambda_{v2}, \ldots, \lambda_{vn}$ be the n eigenvalues of V
Let $\lambda_{z1}, \lambda_{z2}, \ldots, \lambda_{zn}$ be the n eigenvalues of Z
$t\lambda_{v1}, t\lambda_{v2}, \ldots, t\lambda_{vn}$ are the n eigenvalues of tV
Since $tr(A+B) = tr(A) + tr(B)$, we know that the $tr(Z+tV) = tr(Z) + tr(tv) = \lambda_{z1} + t\lambda_{v1} + \lambda_{z2} + t\lambda_{v2} + \cdots + \lambda_{zn} + t\lambda_{vn}$
Let $Z + tV = C$ and let the eigenvalues of C be $\lambda_{c1}, \lambda_{c2}, \ldots, \lambda_{cn}$ be the n eigenvalues of C
We know that $\lambda_{c1} + \cdots + \lambda_{cn} = \lambda_{z1} + t\lambda_{v1} + \lambda_{z2} + t\lambda_{v2} + \cdots + \lambda_{zn} + t\lambda_{vn}$
$g(t) = f(Z+tV) = tr((Z+tV)^{-1}) = \frac{1}{\lambda_{c1}} + \cdots + \frac{1}{\lambda_{cn}}$

$g(t) = f(Z+tV) = tr((Z+tV)^{-1}) = tr(Z^{-1}$
